原始小球藻*AP2*转录因子的生物信息学分析

付春[1,2,*]，唐易[1,2]，江纳[3]，杨瑶君[1,2]

（1. 竹类病虫防控与资源开发四川省重点实验室 乐山师范学院，四川 乐山，614000；2. 乐山师范学院 生命科学学院，四川 乐山，614000；3. 乐山师范学院 数学与信息科学学院，四川 乐山，614000）

摘要：本研究对原始小球藻*AP2*基因家族的蛋白成员进行了详细的生物信息学预测和分析，为进一步研究原始小球藻*AP2*转录因子的功能提供一定的参考依据。利用PrtoParam、Pfam 3.20、Protscale等在线工具分析原始小球藻*AP2*家族所包含的8个蛋白成员。结果表明：该家族成员含有的氨基酸数为247 (XP_011395646.1)到715 (XP_011401904.1)，理论等电点最大为9.39 (XP_011396011.1)、最小为5.81 (XP_011398158.1)；家族中各个成员所含有的保守结构域的位置和数量都各不相同；所有蛋白成员均无信号肽，均不包含跨膜螺旋，不具有跨膜区域；蛋白成员中最大亲水性值为-3.222 (XP_011398158.1)，最大疏水性值为2.333 (XP_011401904.1)，且所有成员的平均亲疏水性值均小于0；成员中有多个磷酸化位点值远超标准值0.5，最大磷酸化位点值为0.998；该基因家族的蛋白成员二级结构的组分含量从大到小排序均为无规则卷曲>α-螺旋>延伸链>β-转角，推测α-螺旋和无规则卷曲是其二级结构的主要方式，延伸链和β-转角则分散在所有蛋白成员的氨基酸链中；所有成员的三级结构分析中均能观察到α-螺旋、β-折叠、β-转角以及N端和C端；系统进化分析结果表明，在其余15个物种中，与原始小球藻亲缘关系最远的是金牛介球菌(*Ostreococcus tauri*)，亲缘关系最为接近的是小球藻(*Chlorella variabilis* NC64A)和螺旋孢子虫(*Helicosporidium*)，较为接近的是苦参3号(*Picochlorum* sp. SENEW3)。该研究较系统地分析了原始小球藻AP2蛋白家族的理化特性、保守基序及系统进化关系，为进一步研究原始小球藻*AP2*转录因子功能和互作关系提供一定的参考依据。

关键词：原始小球藻，*AP2*转录因子，生物信息学

中图分类号：Q944.2　　**文献标识码：**A

Bioinformatics analysis of *AP2* transcription factor in *Auxenochlorella protothecoides*

FU Chun[1,2,*],TANG Yi[1,2],JIANG Na[3],YANG YaoJun[1,2]

(1. *Bamboo Diseases and Pests Control and Resources Development Key Laboratory of Sichuan Province*, *Leshan Normal University*, Leshan 614000, Sichuan, China; 2. *College of Life Science*, *Leshan Normal University*, Leshan 614000, Sichuan, China; 3. *College of Mathematics and Information Science*, *Leshan Normal University*, Leshan 614000 , Sichuan, China)

Abstract： In this study, the protein members of *AP2* gene family in *Auxenochlorella protothecoides* are predicted and analyzed in detail by bioinformatics, which provided some ref-

基金项目：四川省科技厅项目(17ZZ019)；乐山市科技局重点研究项目(18JZD101)；乐山师范学院人才启动项目(XJR17005) [Supported by Sichuan Provincial Science and Technology Department Project (17ZZ019)；Key Research Project of Leshan Municipal Science and Technology Bureau (18JZD101)；Leshan Normal University Talent Launch Project (XJR17005)]。

作者简介：付春(1984-)，男，江西高安人，博士，讲师，主要从事经济植物的生物信息学研究，(E-mail) fuchun421@aliyun.com。

[*]通信作者

erence for further study on the function of *AP2* transcription factor in *Auxenochlorella protothecoides*. Eight protein members of the *AP2* family of *Auxenochlorella protothecoides* are analyzed by online tools such as PrtoParam, Pfam 3.20 and Protscale and so on. The results show that amino acids numbers of the family members are 247(XP_011395646.1) to 715(XP_011401904.1). The maximum theoretical isoelectric point is 9.39(XP_011396011.1), and the minimum is 5.81(XP_011398158.1). The positions and numbers of conserved domains in each member of the family are different; all protein members have no signal peptide; do not contain transmembrane helix, do not have transmembrane region; The maximum hydrophilicity value of the protein members is -3.222(XP_011398158.1), and the maximum hydrophobicity value is 2.333(XP_011401904.1), and the average hydrophilicity/hydrophobicity value of all members is less than 0; The values of many phosphorylation sites in the members were far beyond the standard value of 0.5, and the maximum value of phosphorylation site is 0.998. The component content of secondary structures of protein members in the gene family is arranged in the order of random coil > alpha-helix > extended strand > beta-turn. It is speculated that alpha-helix and random coil are the main modes of its secondary structure, while extended strand and beta-turn are dispersed in the amino acid chain of all protein members.Alpha-helix, beta-sheet, beta-turn, N-terminal and C-terminal were observed in the tertiary structure of all members. Phylogenetic analysis showed that *Ostreococcus tauri* is the most distant relative to the *Auxenochlorella protothecoides*, The closest relatives are *Chlorella variabilis* NC64A and *Helicosporidium*, and the next closest relatives are *Picochlorum* sp. SENEW3 among the other 15 species.This study systematically analyzed the physicochemical properties, conserved motifs and phylogenetic relationships of AP2 protein family in *Auxenochlorella protothecoides*. It provided a reference for further study on the function and interaction of *AP2* transcription factors in *Auxenochlorella protothecoides.*

Key words：*Auxenochlorella protothecoides*, *AP2* transcription factor, bioinformatics

由于化石能源的枯竭，急需在生物能源方面得到突破(侯丽媛，2016)。微藻类生物主要用于生物柴油的研究(夏金兰等，2009)，而原始小球藻(*Auxenochlorella protothecoides*)是一种生长周期短，易培养，可利用光能光合作用进行自养，又能利用有机碳源进行异养培养积累油脂，高附加值产物丰富的微藻(汪桂林等，2013)，它被认为是生物柴油最具潜力的原料之一(桂小华，2017)。原始小球藻生长周期短、适应性强、光合效率和含油量高、单位面积产量高，现已被作为生物柴油方面主要能源藻种加以研究开发(汪桂林等，2013)。目前对藻类的*AP2*转录因子研究较少。

*AP2*是植物所特有的转录因子家族之一，起到调控植物生长发育的转录以及逆境胁迫等生理过程的作用(栗丽和佟少明，2016)。1994年Jofuku(Jofuku et al., 1994)等从模式植物拟南芥(*Arabidopsis thaliana*)中分离了第一个*AP2*基因，且该基因与花发育有关(张计育等，2012)，其后又在原核生物和蓝藻中发现了相同的转录因子(张麒等，2018；Wessler, 2005)，证实该转录因子不止只存在于植物中。在不同的植物中含有的*AP2/ERF* 类转录因子基因家族成员数量是不同的(Riccardo Velasco et al., 2010)，转录因子*AP2/ERF*主要特征是具有高度保守的约60~70个氨基酸组成的AP2结构域(闻可心和刘雪梅，2010)，这是植物转录因子家族之一，家族成员均包含保守AP2结构域，人们根据保守结构域的数量和识别成员的不同，通常情况下，将*AP2*/*ERF*家族划分为不同的4个转录因子亚家族，分别是：*AP2*，*ERF*，*DREBP*，*RAV*和单独成员Soloist(Zhuang et al., 2008；张麒等，2018)。转录因子*AP2*基因家族在植物生长发育中起很大的作用，它能调控植物的细胞周期、生长发育以及生物和非生物胁迫相关基

因的表达(庄静等，2008；陈茜，2011；吴丽娟等，2016；郭慧等，2017)，也能调控自身生殖器官(花、胚珠和种子)的生长发育(赵利锋和柴团耀，2008；张妍和孙丰宾，2011；王莉莉，2016)，而且与其他因子相互协作，参与到复杂的花发育调控网络(刘杨，2009)。虽然关于藻类*AP2*基因家族的研究比较多，但是未见关于原始小球藻*AP2*转录因子的生物信息学分析的报道。

本研究通过分析原始小球藻的*AP2*转录因子编码蛋白的理化特性、亲疏水性、跨膜域、保守结构域、保守基序、亚细胞定位以及*AP2*转录因子在不同物种中的进化起源关系，为进一步研究原始小球藻*AP2*基因家族的生物学功能奠定一定的理论依据。

1 材料与方法

1.1 材料

从plantTFDB (http：//planttfdb.cbi.pku.edu.cn/)数据库中下载原始小球藻(*Auxenochlorella protothecoides*)*AP2*基因家族中所有成员的蛋白序列和岩藻(*Bathycoccus prasinos*)、莱茵衣藻(*Chlamydomonas reinhardtii*)、小球藻(*Chlorella variabilis* NC64A)、胶球藻(*Coccomyxa subellipsoidea* C-169)、盐生杜氏藻(*Dunaliella salina*)、胸状盘藻(*Gonium pectorale*)、螺旋孢子虫(*Helicosporidium*)、细小微胞藻(*Micromonas pusilla* CCMP1545)、微单胞藻(*Micromonas* sp. RCC299)、深蓝单壳缝藻(*Monoraphidium neglectum*)、透明介球菌(*Ostreococcus lucimarinus*)、介球菌(*Ostreococcus* sp. RCC809)、金牛介球菌(*Ostreococcus tauri*)、苦参3号(*Picochlorum* sp. SENEW3)以及团藻(*Volvox carteri*)等物种的AP2成员蛋白序列。对原始小球藻的AP2所有成员的蛋白质序列进行详细的生物信息学分析，且将原始小球藻与岩藻(*Bathycoccus prasinos*)等16个物种AP2蛋白进行同源序列比对分析并构建进化树，探究16个物种间的进化起源关系。

1.2 方法

1.2.1 *AP2*转录因子基因的蛋白成员的一级结构和理化性质分析

利用在线程序PrtoParam (https://web.expasy.org/protparam/)对原始小球藻*AP2*转录因子蛋白成员的一级结构和理化性质进行分析，包括了蛋白成员氨基酸数目、理论pI、分子量、带负电氨基酸残基、带正电氨基酸残基、不稳定系数、脂肪指数、亲疏水性、分子式、原子总数等方面(韩长志和王娟，2016)。

通过在线程序Cell-PLoc 2.0 (http://www.csbio.sjtu.edu.cn/bioinf/ Cell-PLoc-2/)对原始小球藻*AP2*转录因子的8个蛋白成员进行亚细胞定位点分析(徐志军等，2019)，参数设置：物种选择为真核生物。

1.2.2 氨基酸的保守结构域、跨膜结构及信号肽分析

利用Pfam 3.20 (http://pfam.xfam.org/)对原始小球藻的蛋白成员进行保守结构域预测(张海燕等，2013)，统计分析蛋白成员氨基酸保守结构域的个数及位置。

在TMHMM Severv.2.0 (http://www.cbs.dtu.dk/ services/TMHMM/)在线程序中输入*AP2*转录因子基因蛋白成员，分析跨膜螺旋的数量以及跨膜结构的区域，确定基因蛋白成员是否具有跨膜区域，判断是否为跨膜蛋白。

利用SignaIP3.0 Server(http://www.cbs.dtu.dk/services/SignalP-3.0/)对原始小球藻蛋白成员进行信号肽分析(柳燕等，2017)。

1.2.3 氨基酸的亲疏水性和磷酸化位点预测分析

利用Protscale (https://web.expasy.org/protscale/)在线工具，对该转录因子基因蛋白成员进行亲疏水性预测分析，确定最大亲水值和最大疏水值及其相对应的位置，判断该基因蛋白

成员为亲水蛋白还是疏水蛋白。

使用NetPhos 3.1 Server (http://www.cbs.dtu.dk/services/NetPhos/)在线程序对转录因子基因蛋白成员进行潜在磷酸化位点预测分析，并统计磷酸化位点种类和数量。

1.2.4 编码蛋白的二级结构和三级结构预测分析

利用SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html)在线程序对该转录因子蛋白成员的二级结构进行预测分析，统计该蛋白成员中α-螺旋、无规则卷曲、延伸链和β-转角相对应的百分比。

通过同源建模法，利用SWISS-MODEL (https://www.swissmodel.expasy.org/)在线工具(Schwede T, et al., 2003)，对该蛋白家族的三级结构进行预测分析，确定其三级结构的组成和构象变化。

1.2.5 保守基序分析

利用MEME7.0 (http://meme-suite.org/)在线工具对原始小球藻*AP2*转录因子的蛋白成员进行保守基序motif分析，参数设置：motif数目为10个，其余参数设为默认值。

1.2.6 同源性分析及进化树分析

在 plantTFDB (https://web.expasy.org/protparam/)在线程序中，将原始小球藻的 8 个成员序列，分别和同为绿藻门下的其余 15 个物种 *AP2* 基因家族的蛋白成员进行同源序列的对比，找到相对应的其余物种的同源序列，并归纳。

利用 MEGA7.0 (Tamura K, et al., 2011)软件中的 ClustalW 功能对原始小球藻 AP2 转录因子所有成员蛋白对应的序列组，分别进行多重比对并邻接法构建其系统进化树(李建伏等, 2008)，对比 8 个蛋白成员的同源序列的差距，判断各个物种与原始小球藻之间的亲缘关系。

2 结果与分析

2.1 *AP2*转录因子基因成员的一级结构和理化性质分析

原始小球藻*AP2*转录因子蛋白家族含有8个成员，该蛋白家族氨基酸数目范围为247~715；等电点范围为5.81(XP_011398158.1)~9.39(XP_011396011.1)，理论等电位点>7为碱性氨基酸：XP_011395646.1、XP_011396011.1、XP_011400319.1和XP_011401904.1，理论等电位点<7为酸性氨基酸：XP_011396127.1、XP_011401790.1、XP_011400550.1和XP_011398158.1；该家族成员带负电氨基酸残基数目范围为26(XP_011395646.1)~85(XP_011401790.1)；该家族成员带正点氨基酸残基数目范围为26(XP_011398158.1)~81(XP_011401790.1)；不稳定指数均大于40(不稳定系数<40时稳定)(夏巧玉，2007)，推断为不稳定蛋白；脂肪指数范围为58.87(XP_011401790.1)~74.26(XP_011400550.1)；原始小球藻AP2蛋白成员进行的亚细胞定位分析中显示该家族全部成员的亚细胞都可定位于细胞质(Cytoplasm)，其中4个蛋白成员成员：XP_011396011.1、XP_011400319.1、XP_011401790.1和XP_011398158.1的亚细胞只定位于细胞质(表1)。通过氨基酸组成分析可知，在所有蛋白成员中，除了XP_011396011.1(氨基酸的最大含量为甘氨酸-Gly-15.80%)和XP_011398158.1(氨基酸的最大含量为亮氨酸-Leu-11.10%)两个蛋白家族成员外，其余六个蛋白家族成员含量最高的氨基酸均为丙氨酸(Ala)，丙氨酸含量最高的成员为XP_011400550.1(Ala19.70%)，含有丙氨酸数为60，丙氨酸含量最低的成员为XP_011400319.1(Ala11.2%)，含有丙氨酸数为38(图1)。

表1 原始小球藻AP2蛋白的理化特性

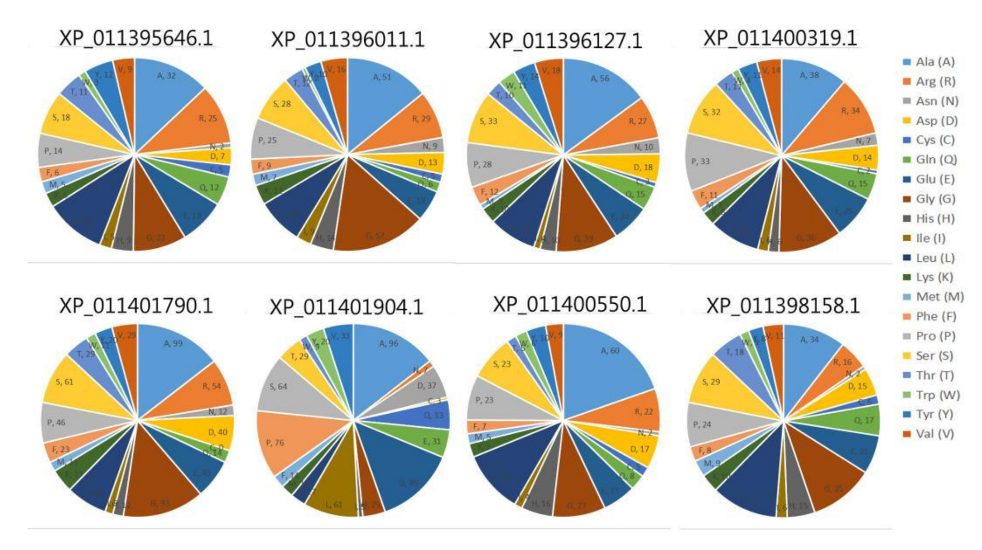
Table 1 Physical and chemical properties of AP2 protein in *Auxenochlorella protothecoides*

蛋白ID Protein ID	氨基酸数 Number of amino acids	分子量 Molecular	理化特性 Physicochemical property					
			理论pI	带负电	带正电	不稳定指数	脂肪指数	亚细胞定位

		weight	Theoretical pI	残基 (Asp+Glu)	残基 (Arg+Lys)	Instability index	Aliphatic index	Subcellular localization
XP_011395646.1	247	27377.91	8.98	26	31	66.10	70.89	细胞质，细胞外 Cytoplasm, extracell
XP_011396011.1	361	37242.76	9.39	32	41	46.47	66.95	细胞质 Cytoplasm
XP_011396127.1	375	40364.66	6.08	42	37	49.28	64.21	细胞质，细胞外 Cytoplasm, extracell
XP_011400319.1	339	36701.84	8.83	39	42	74.47	63.45	细胞质 Cytoplasm
XP_011401790.1	682	71540.98	6.37	85	81	42.67	58.87	细胞质 Cytoplasm
XP_011401904.1	715	74708.18	7.86	68	69	58.52	62.81	细胞质，细胞核 Cytoplasm, nucleus
XP_011400550.1	305	32414.34	6.34	34	29	55.92	74.26	细胞质，细胞外 Cytoplasm, extracell
XP_011398158.1	324	34742.98	5.81	36	26	47.43	70.9	细胞质 Cytoplasm

注: (Asp+Glu)为带负电荷的残基总数；(Arg+Lys)为带正电荷的残基总数。

Note: (Asp+Glu) is the total of the residue of the negatively charged; (Arg+Lys) is the total of the residue of the positive charged.



A.丙氨酸；**R**. 精氨酸；**N**. 天冬酰胺；**D**. 天冬氨酸；**C**. 半胱氨酸；**Q**. 谷氨酰胺；**E**. 谷氨酸；**G**. 甘氨酸；**H**. 组氨酸；**I**. 异亮氨酸；**L**. 亮氨酸；**K**. 赖氨酸；**M**. 蛋氨酸；**F**. 苯丙氨酸；**P**. 脯氨酸；**S**. 丝氨酸；**T**. 苏氨酸；**W**. 色氨酸；**Y**. 酪氨酸；**V**. 缬氨酸。

A. Alanine; **R**. Arginine; **N**. Apsaragine; **D**. Apsartic acid; **C**. Cystine; **Q**. Glutarnine; **E**. Glutamicacid; **G**. Glycine; **H**. Histidine; **I**. Isoleucine; **L**. Leucine; **K**. Lysine; **M**. Methionine; **F**. Phenylalanine; **P**. Proline; **S**. Serine; **T**. Threonine; **W**. Tryptophan; **Y**. Tyrosine; **V**. Valine.

图1 原始小球藻AP2蛋白质的氨基酸组成

Fig.1 Amino acid composition of AP2 protein from *Auxenochlorella protothecoides*

2.2氨基酸保守结构域、信号肽和跨膜域结构分析

蛋白成员中，不同基因蛋白成员的AP2保守结构域的数量和位置各不相同，其中XP_011396127.1(K_{38}-L_{96},R_{118}-Y_{178},F_{200}-E_{260}) 和 XP_011401790.1(S_{67}-D_{104},W_{188}-A_{233},S_{264}-E_{322})两个成员各含3个保守结构域；而XP_011398158.1(Y_{50}-E_{111})和XP_011400550.1(Y_{29}-E_{96})两个成员中各自都只含有1个氨基酸保守结构域；余下的XP_011395646.1(Y_{59}-T_{111},F_{151}-A_{207})、XP_011396011.1(Q_{169}-A_{233},S_{264}-G_{322})、XP_011400319.1(K_{174}-E_{235},Y_{265}-E_{326})和XP_011401904.1(Q_{258}-E_{333},F_{363}-E_{426})4个成员，均包含2个保守结构域。保守结构域跨度最大成员是XP_011401904.1(Q_{258}-E_{333})，其氨基酸数量达75个，蛋白成员XP_011401790.1(S_{67}-D_{104})中保守结构域跨度最小，其氨基酸数量为37个(表2，图2)。各个蛋白成员的位置和跨度都各不相同，但却有相似，如XP_011395646.1、XP_011400319.1和XP_011401904.1这3个成员就含有相似之处(图2)。信号肽分析表明原始小球藻AP2蛋白均不存在信号肽，属于非分泌型蛋白。TMHMM跨膜域结构分析表明原始小球藻所有蛋白成员均不含跨膜结构域，推断该蛋白为非跨膜蛋白。

表2 原始小球藻AP2蛋白保守结构域数量和位置

Table 2 Number and location of conserved domains of AP2 protein in *Auxenochlorella protothecoides*

蛋白ID Protein ID	AP2保守结构域个数 Number of conservative structural domains in AP2	保守结构域的位置 Conservative domain location
XP_011395646.1	2	Y_{59}-T_{111},F_{151}-A_{207}
XP_011396011.1	2	Q_{169}-A_{233},S_{264}-G_{322}
XP_011396127.1	3	K_{38}-L_{96},R_{118}-Y_{178},F_{200}-E_{260}
XP_011400319.1	2	K_{174}-E_{235},Y_{265}-E_{326}
XP_011401790.1	3	S_{67}-D_{104},W_{188}-A_{233},S_{264}-E_{322}
XP_011401904.1	2	Q_{258}-E_{333},F_{363}-E_{426}
XP_011400550.1	1	Y_{29}-E_{96}
XP_011398158.1	1	Y_{50}-E_{111}



图2 原始小球藻AP2蛋白氨基酸保守域预测

Fig.2 Prediction of amino acids conservative domain of AP2 protein in *Auxenochlorella protothecoides*

2.3 氨基酸亲疏水性预测分析

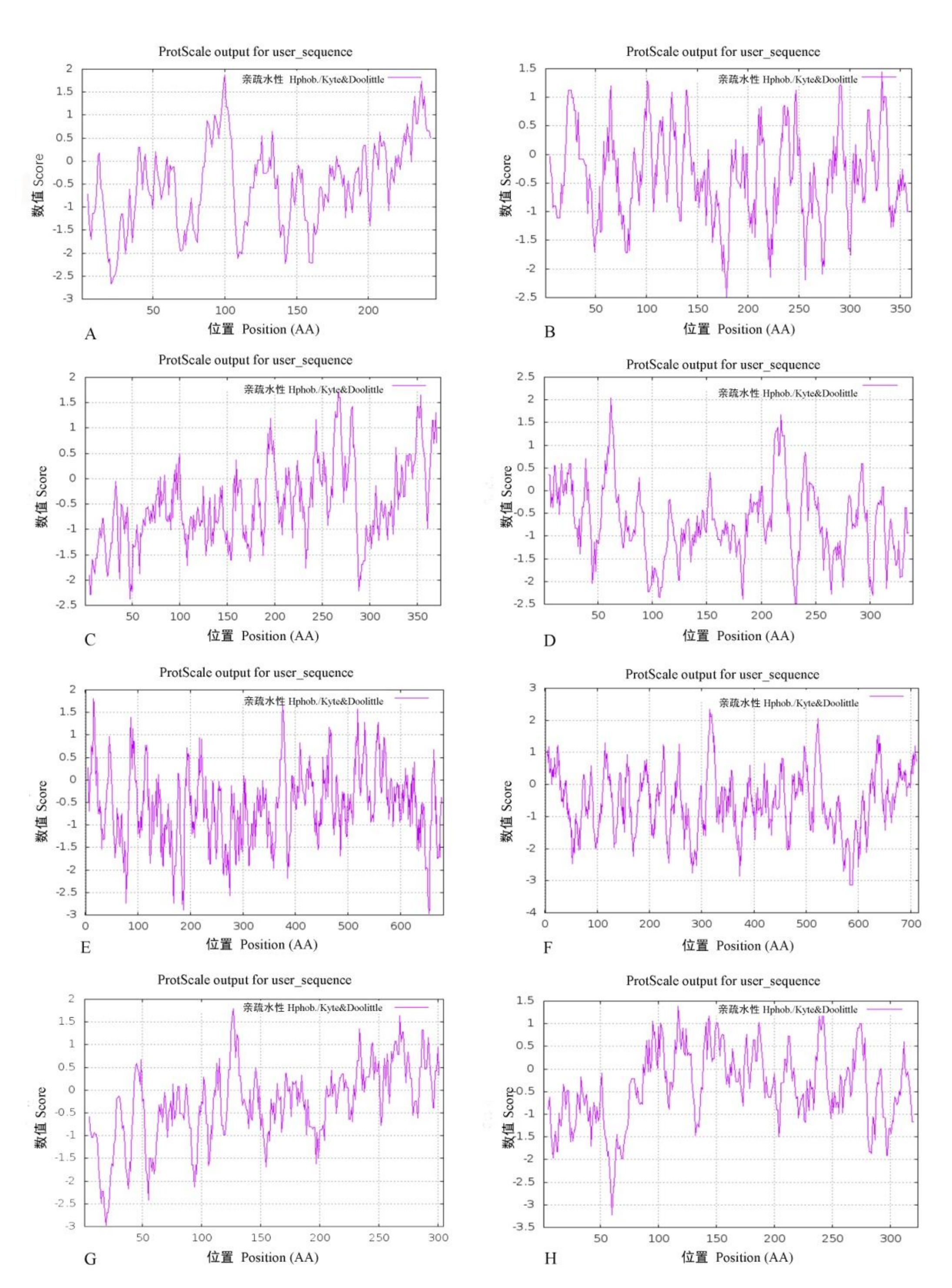
在这8个蛋白家族成员中，氨基酸疏水性的最大值达到了2.333 (XP_011401904.1)，疏水性最小值为1.389 (XP_011398158.1)；氨基酸亲水性最大值达到了-3.222 (XP_011398158.1)，亲水性最小值为-2.378 (XP_011396127.1)。该家族成员XP_011401904.1的氨基酸亲疏水性差值最大，达到了5.455，对应平均亲水指数(GRAVY)为-0.3945，其第316位氨基酸表现最大疏水性为2.333，第584、585、586、588、589位氨基酸均表现最大亲水性为3.122。所有的AP2蛋白家族成员中的平均亲水指数均小于0，所以该蛋白亲水性更强，推断该原始小球藻*AP2*转录因子蛋白为亲水性蛋白(表3，图3)。坐标图中，纵坐标显示氨基酸标度值，正值表示疏水，负值表示亲水，横坐标显示氨基酸序列位置(图3) (杨朗，2015) 。

表3 原始小球藻AP2蛋白的亲疏水性

Table 3 Hydrophilicity of AP2 protein from *Auxenochlorella protothecoides*

蛋白ID Protein ID	亲水性 Hydrophilicity		疏水性 Hydrophobicity	
	数值 Value	位置 Position	数值 Value	位置 Position
XP_011395646.1	-2.667	21	1.878	100
XP_011396011.1	-2.500	179	1.444	332
XP_011396127.1	-2.378	48	1.822	267
XP_011400319.1	-2.489	231	2.044	62
XP_011401790.1	-2.967	654	1.800	16
XP_011401904.1	-3.122	586	2.333	316

XP_011400550.1	-2.967	19	1.789	127
XP_011398158.1	-3.222	60	1.389	117



A. XP_011395646.1; **B**. XP_011396011.1; **C**. XP_011396127.1; **D**. XP_011400319.1; **E**. XP_011401790.1; **F**. XP_011401904.1; **G**. XP_011400550.1; **H**. XP_011398158.1.

图3 原始小球藻AP2蛋白家族的亲水性分析

Fig.3 Hydrophilicity analysis of AP2 protein family of *Auxenochlorella protothecoides*

2.4 磷酸化位点分析

经分析发现，在阈值为0.5时,原始小球藻AP2转录因子蛋白成员中潜在的磷酸化位点最

多的是XP_011401790.1 (77)，潜在磷酸化位点最少的是XP_011400550.1 (24)。将同种氨基酸对比分析，在整个家族中，潜在位点为丝氨酸(S)数量最多的是XP_011401790.1，达到了49个位点,数量最少的是XP_011395646.1和XP_011400550.1两个成员，均含有17个位点；潜在位点为苏氨酸(T)数量最多的是XP_011401790.1，达到了24个位点，数量最少的是XP_011400550.1，仅含有4个位点；潜在位点为络氨酸(Y)数量最多的是XP_011396011.1和XP_011401904.1，都达到了5，数量最少的是XP_011398158.1，仅1个位点(表4)。在第4个分支XP_011400319.1中，第164位、第218位的苏氨酸，第142位、第160位丝氨酸以及第268位络氨酸可能是潜在的磷酸化位点，其中最有可能是潜在磷酸化位点的是第142位和第160位的丝氨酸，其值均已经达到了0.998，远超了标准值0.5，同时也说明蛋白质可能通过丝氨酸磷酸化位点来实现其功能的调控(表4)。

表4 原始小球藻AP2蛋白家族的磷酸化位点分析

Table 4 Phosphorylation site analysis of AP2 protein family in *Auxenochlorella protothecoides*

蛋白ID Protein ID	氨基酸 Amino acid	个数 Number	值 Value	位置 Position
XP_011395646.1	丝氨酸 Serine	17	0.998	59,60,72
	苏氨酸 Threonine	10	0.984	86
	络氨酸 Tyrosine	2	0.877	148
XP_011396011.1	丝氨酸 Serine	21	0.997	214,290
	苏氨酸 Threonine	8	0.844	143
	络氨酸 Tyrosine	5	0.978	119
XP_011396127.1	丝氨酸 Serine	26	0.997	239
	苏氨酸 Threonine	8	0.991	52
	络氨酸 Tyrosine	4	0.983	339
XP_011400319.1	丝氨酸 Serine	29	0.998	142,160
	苏氨酸 Threonine	9	0.989	164,218
	络氨酸 Tyrosine	3	0.705	268
XP_011401790.1	丝氨酸 Serine	49	0.998	305
	苏氨酸 Threonine	24	0.965	683
	络氨酸 Tyrosine	4	0.951	677
XP_011401904.1	丝氨酸 Serine	47	0.997	102,389
	苏氨酸 Threonine	19	0.995	205,241
	络氨酸 Tyrosine	5	0.895	366
XP_011400550.1	丝氨酸 Serine	17	0.993	59,204,277
	苏氨酸 Threonine	4	0.664	45
	络氨酸 Tyrosine	3	0.984	217
XP_011398158.1	丝氨酸 Serine	18	0.969	239
	苏氨酸 Threonine	11	0.954	64
	络氨酸 Tyrosine	1	0.514	166

2.5 原始小球藻*AP2*转录因子编码蛋白的二级结构和三级结构预测分析

蛋白质的二级结构主要是指蛋白质的多肽链中有规则重复的构象（曹晨，2016），在原始小球藻 AP2 转录因子基因蛋白成员中，二级结构中组分含量从大到小排序均为无规则卷曲>α-螺旋>延伸链>β-转角，其中无规则卷曲和 α-螺旋的占比较大，推测其空间结构较小，

且受到侧链间相互作用较大的影响(王月，2017)。基因二级结构组分中α-螺旋(Alpha helix)值最大的是 XP_011400550.1 成员成员，占比达到了 40.00%，α-螺旋值最小的是 XP_011396011.1 成员成员，其占比为 26.04%；该蛋白三级结构组分中延伸链占比最大为 XP_011400319.1，达 16.22%，其延伸链最小占比为 XP_011396127.1，达 10.40%；该组分中 β-转角值为 3.64% (XP_011395646.1)~8.64% (XP_011398158.1)；其结构组分中无规则卷曲值为 40.00% (XP_011400550.1)~54.55% (XP_011401904.1)(表 6)。由此可推断，在原始小球藻 AP2 基因蛋白成员中 α-螺旋和无规则卷曲是二级结构的主要方式，延伸链和 β-转角则分散在整个基因蛋白成员的氨基酸链中(表 5，图 4)。

表5 原始小球藻AP2蛋白质二级结构元件含量占比统计

Table 5 Statistical analysis of the content of secondary structural elements of AP2 protein in *Auxenochlorella protothecoides*

蛋白ID Protein ID	α-螺旋 Alpha helix (Hh)	延伸链 Extended strand (Ee)	β-转角 Beta turn (Tt)	无规则卷曲 Random coil (Cc)
XP_011395646.1	36.03%	12.96%	3.64%	47.37%
XP_011396011.1	26.04%	14.68%	8.03%	51.25%
XP_011396127.1	30.40%	10.40%	5.07%	54.13%
XP_011400319.1	26.55%	16.22%	5.31%	51.92%
XP_011401790.1	29.33%	11.14%	8.21%	51.32%
XP_011401904.1	26.15%	10.91%	8.39%	54.55%
XP_011400550.1	40.00%	11.80%	8.20%	40.00%
XP_011398158.1	35.80%	10.49%	8.64%	45.06%

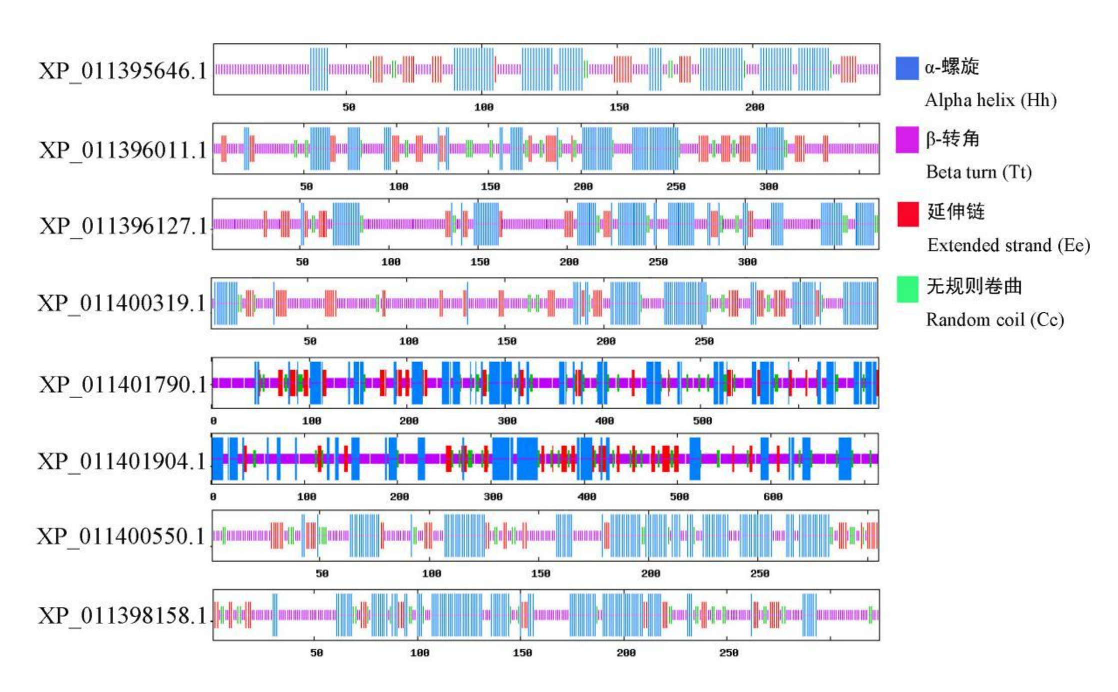


图4 原始小球藻AP2蛋白质二级结构预测分析

Fig.4 Prediction and analysis of secondary structure of AP2 protein in *Auxenochlorella protothecoides*

蛋白质三级结构预测分析表明原始小球藻AP2家族蛋白成员均含有α-螺旋(α-helix)、β-折叠(β-sheet)、β-转角(β-turn)以及成员的N端(N-terminal)和C端(C-terminal)等结构，其中最主要的结构是α-螺旋(α-helix)，其次是β-折叠(β-sheet)。且XP_011398158.1、XP_011400550.1

和XP_011398158.1三个成员的三级结构高度相似，而XP_011400319.1和XP_011400319.1两个成员的三级结构也同样高度相似(图5)。

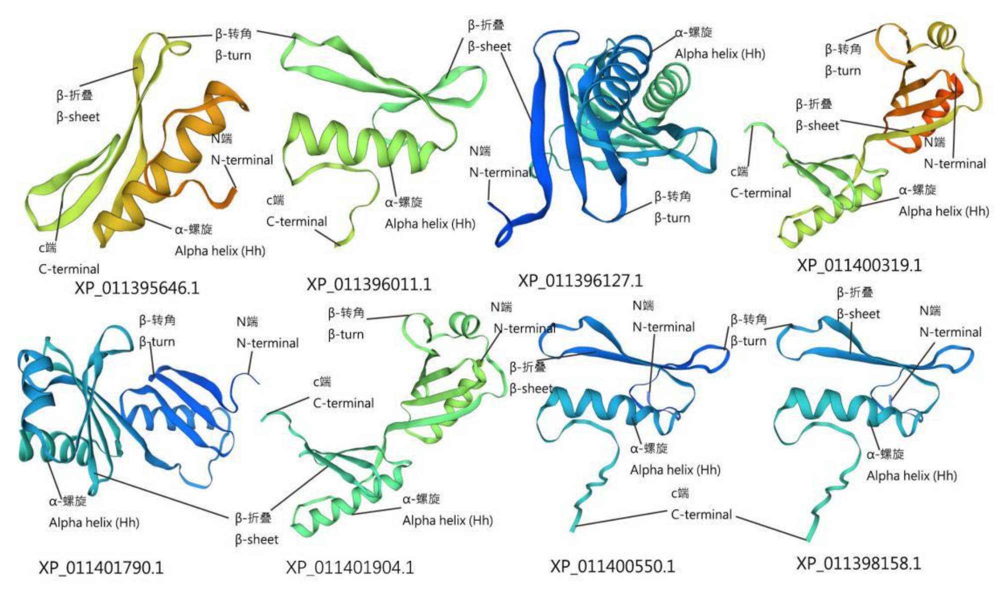


图5 原始小球藻AP2蛋白的三级结构预测

Fig.5 Prediction of the tertiary structure of AP2 protein from *Auxenochlorella protothecoides*

2.6 保守基序MEME分析

通过分析可知，在原始小球藻*AP2*转录因子家族的8个成员中，motif1只存在于XP_011395646.1和XP_011401904.1两个成员中；除XP_011400550.1和XP_011398158.1两个成员外，其余6个成员中均含有motif2和motif4；所有成员均包含motif3和motif5；motif6存在于XP_011395646.1和XP_011396011.1两个成员中；motif7存在于XP_011396127.1和XP_011398158.1两个成员中；motif8存在于XP_011396127.1和XP_011401790.1两个成员中；motif9存在于XP_011400319.1、XP_011396127.1和XP_011398158.1等3个成员中。该家族成员可分类为4个，其中XP_011395646.1和XP_011396011.1为一类，都含有的是motif2、motif3,、motif4、motif5、motif6和motif10；第二类是XP_011400319.1和XP_011401904.1，都含有motif3、motif4和motif5；第三类是XP_011396127.1和XP_011401790.1，都含有motif2、motif3、motif4和motif8；第四类是XP_011400550.1和XP_011398158.1，都含有motif3和motif5。在原始小球藻AP2蛋白成员中，最主要含有的是motif3和motif5, motif3和motif5可能是参与植物细胞生长发育和逆境胁迫等多个生理过程有关的功能基序(栗丽和佟少明，2016)(图6，图7)。

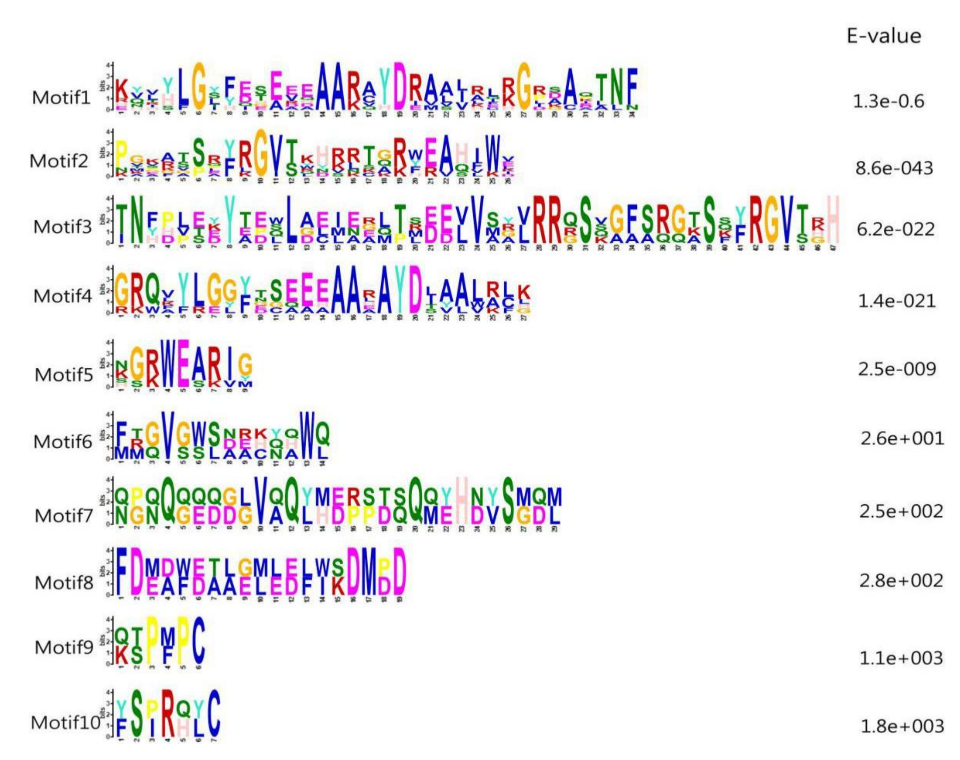


图6 原始小球藻AP2家族蛋白成员保守基序分析

Fig.6 Conserved motif analysis of AP2 family proteins in *Auxenochlorella protothecoides*

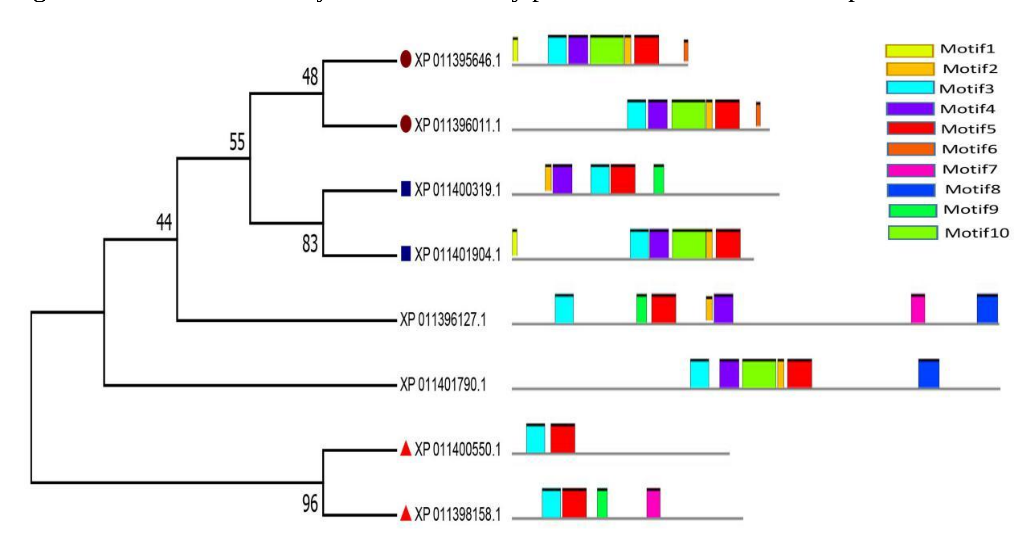


图7 原始小球藻*AP2*蛋白的保守基序分布

Fig.7 Distribution of conserved motifs of AP2 protein in *Auxenochlorella protothecoides*

2.7同源性比对及进化树分析

将原始小球藻的8个成员序列分别在plantTFDB在线程序中，和与原始小球藻同为绿藻门下的其余15个物种*AP2*基因家族的蛋白成员进行同源序列的对比，找到相对应的其余物种的同源序列，并分别使用MEGA7.0软件对8个蛋白成员序列组进行系统进化树的构建，对比8个蛋白成员的同源序列的差距，判断各个物种与原始小球藻之间的亲缘关系。在XP_011396127.1和XP_011396011.1两个进化树中亲缘关系最远的是金牛介球菌，在XP_011395646.1进化树中最远的是团藻，在XP_011400319.1进化树中，最远的是盐生杜氏藻，在XP_011401790.1进化树中，最远的是莱茵衣藻，在XP_011401904.1进化树中，最远的是胶球藻，在XP_011400550.1进化树中，最远的是深蓝单壳缝藻，在XP_011398158.1进

化树中，最远的是微单胞藻，根据此推断，16个物种中与原始小球藻亲缘关系最远的是金牛介球菌；在XP_011395646.1、XP_011396011.1、XP_011400550.1和XP_011401904.1这4个进化树中都是螺旋孢子虫与原始小球藻的亲缘关系最近，在XP_011396127.1和XP_011398158.1进化树中，亲缘关系最近的是小球藻，在XP_011401790.1进化树中，亲缘关系最近的是胶球藻，在XP_011400319.1进化树中亲缘关系最近的是苦参3号；而在所有的进化树中，小球藻与原始小球藻亲缘关系在同一大分支中一共有7次，苦参3号有6次，螺旋孢子虫有7次，胶球藻有3次，盐生杜氏藻有1次，莱茵衣藻有2次，深蓝单壳缝藻、胸状盘藻和团藻各有1次，根据此推断，16个物种中与原始小球藻亲缘关系最为接近的是小球藻和螺旋孢子虫，较为接近的是苦参3号 (图8) 。

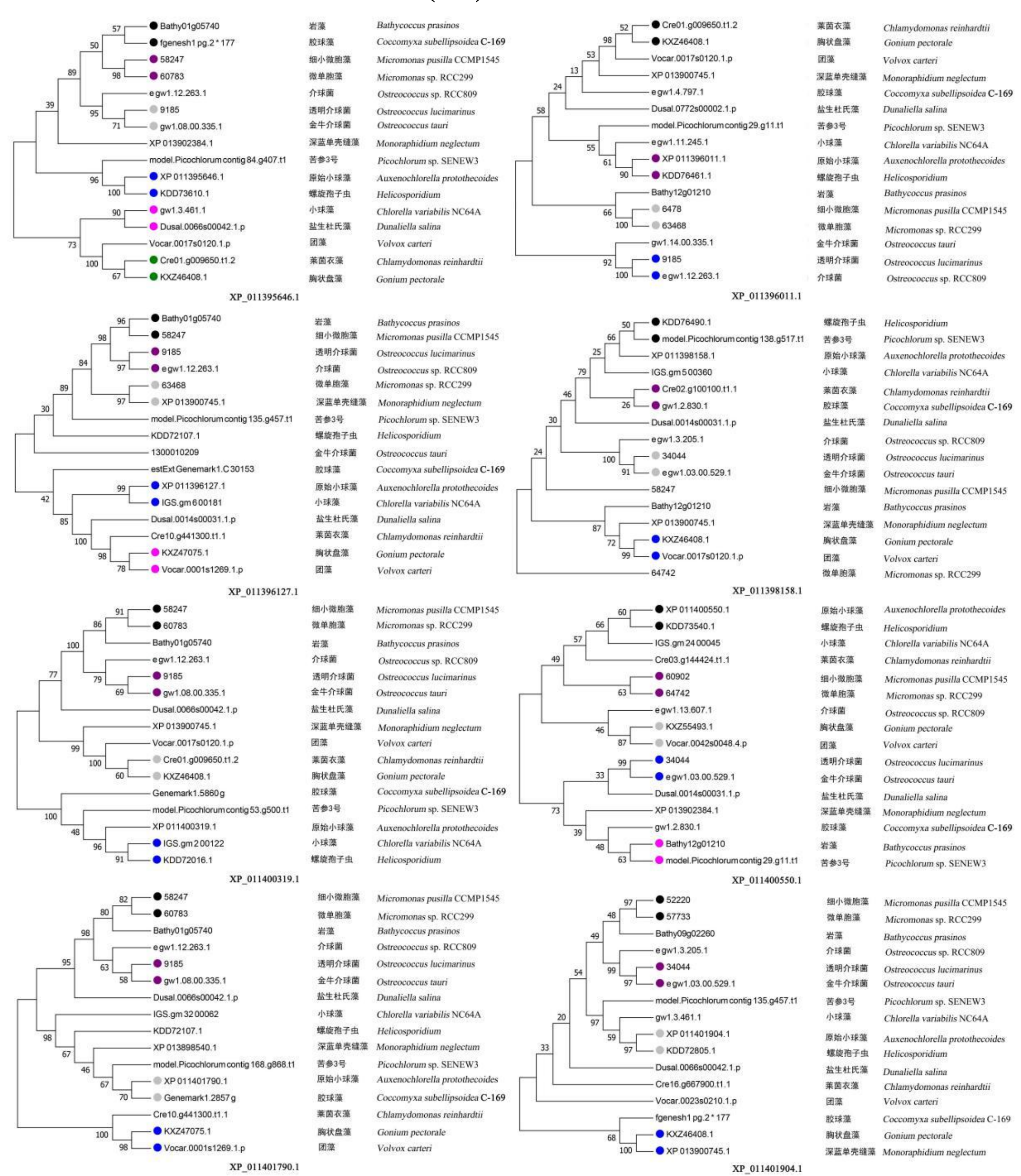


图8 原始小球藻AP2蛋白家族的系统发育树分析

Fig.8 Phylogenetic tree analysis of AP2 protein family in *Auxenochlorella protothecoides*

3 讨论

通过对原始小球藻*AP2*基因蛋白成员进行详细的生物信息学分析后，可较详细的了解原

始小球藻*AP2*基因蛋白成员的各种特性。原始小球藻中含有8个家族成员，而在拟南芥(*Arabidopsis*)中为18个(唐美芳，2007)，水稻(*Oryza sativa*)中为26个(唐美芳，2007；施杨等，2013)，玉米(maize)中为31个(莫晓婷等，2013)，由此可见不同物种的AP2家族蛋白成员数量是不一样的。原始小球藻属于植物界、小球藻科、真核生物(吴丽娟等，2016)、在系统进化中最为原始，该蛋白家族成员含有247到715个氨基酸，在其*AP2*转录因子蛋白成员中，含有多个保守结构域，但是不存在信号肽和跨膜结构域，且AP2蛋白的亲和性平均值均小于0，故其为非分泌、非跨膜、亲水性型蛋白。在原始小球藻的理化性质分析中，*AP2*家族成员共编码3348个氨基酸，其中同时存在碱性氨基酸和酸性氨基酸，不稳定系数>40，为不稳定蛋白，而在雷公藤(*Tripterygium wilfordii*)(祝传书等，2018)、茶树(*Camellia sinensis*)(吴致君等，2014)和枣(*Ziziphus jujube*)(纪晴等，2018)中的编码氨基酸同样含有碱性氨基酸和酸性氨基酸、为不稳定蛋白，但茶树含有的碱性氨基酸略多于酸性氨基酸，且蛋白质残基为210~290；而枣的酸性氨基酸多于碱性氨基酸，这与本文研究结果相似。在保守结构域的分析中，不同物种的个数及位置不尽相同：原始小球藻和枣(*Ziziphus jujube*)(纪晴等，2018)的AP2成员中同样存在3个保守结构域，其中原始小球藻的氨基酸数量跨度最大达到75，但野生大豆(*Glycine soja*)(朱延明等，2019)的保守结构域个数只存在两个，其氨基酸数量跨度较大达到104(209~313)。在亲疏水性分析中，原始小球藻AP2成员蛋白与菠萝 AcPEPC 蛋白(马海洋等，2020)和蔓花生*PEPC*基因家族(涂嘉琦等，2018)相同，所有蛋白成员均为亲水性蛋白，磷酸化位点分析中，原始小球藻AP2蛋白家族成员含有多个潜在磷酸化位点，最有可能的磷酸化位点为丝氨酸，在蓖麻(*Ricinus communis*)APs分析中潜在磷酸化位点同样为丝氨酸。在原始小球藻的二级结构分析中，其二级结构的组分含量从大到小排序均为无规则卷曲>α-螺旋>延伸链>β-转角，α-螺旋和无规则卷曲是二级结构的主要方式，延伸链和β-转角则分散在所有蛋白成员的氨基酸链中：在中国樱桃(Chinese cherry)(王月，2017)和甘薯(*Dioscorea esculenta*)(阮先乐等，2017)中的二级结构占比分布和原始小球藻相同，α-螺旋和无规则卷曲占比最大；但在杨树(*Pterocarya stenoptera*)(赵金玲等，2015)中，其组分与原始小球藻有所差异，仅以无规则卷曲为主。在三级结构分析中，甘蓝与原始小球藻都能明显观察到的是α-螺旋(α-helix)和β-折叠(β-sheet)，但原始小球藻三级结构家族成员之间结构差异较大，而甘蓝(*Brassica oleracea*)(郭慧等，2017)和甘薯(*Dioscorea esculenta*)(阮先乐等，2017)的家族成员三级结构基本相似。在原始小球藻保守基序中，将所有基序一共分为4类，其中所有成员共同含有的是motif3和motif5，motif3和motif5可能是参与原始小球藻细胞生长发育和逆境胁迫等多个生理过程有关的功能基序(栗丽和佟少明，2016)。在原始小球藻AP2蛋白家族中不存在信号肽和跨膜结构域，同样在沙棘(*Hippophae rhamnoides*)*WRI1* 转录因子(马倩等，2016)蛋白序列和蔷薇科植物的19个DELLA蛋白(宋伟等，2013)的序列中也均无信号肽和跨膜结构域，但同时也存在不同，例如：在花生profilin蛋白(肖杰等，2011)序列中，存在信号肽位于第1~24位蛋白，却没有跨膜结构域；在大豆GmANKTM家族(柏锡等，2019)蛋白中存在跨膜结构域；在小麦(*Triticum aestivum*)F-box蛋白(王俊生等，2013)基因序列中，第94至第115位存在一个可能的跨膜区，且在该蛋白质N端序列不存在信号肽

目前在植物界对*AP2*转录因子的研究主要在单双子叶植物上，如：在拟南芥(*Arabidopsis*)(Jofuku et al.，1994)中发现124个*AP2/ERF*家族基因，水稻(*Oryza sativa*)(Nakano, 2006)中发现了139个*AP2/ERF*家族基因，大豆(*Glycine max*)中发现了120个*AP2/ERF*家族基因(Zhang et al., 2008)，小兰屿蝴蝶兰(*Phalaenopsis equestris*)中发现了107个*AP2/ERF*家族基因等(朱光哲，2016)，而在藻类上*AP2*转录因子的研究较少。本研究所主要描述的是对原始小球藻*AP2*所有蛋白成员进行了详细的生物信息学分析，系统进化分析结果说明了原始小球藻最为原始，与原始小球藻亲缘关系最远的是细小微胞藻ccmp1545 (*Micromonas pusilla* CCMP1545)，亲缘关系最近的是团藻(*Volvox carteri*)。本研究结果为今后进一步研究*AP2*转

录因子在原始小球藻生长发育过程中的生理特点和分子功能提供了一定的参考依据。

参考文献：

BO X, WANG XY, CHEN Y, et al., 2019. Bioinformatics analysis of GmANKTM family abiotic stress in soybean[J]. J NE Agric Univ, 50(7):18-27. [柏锡，王昕奕，陈云，等，2019. 大豆GmANKTM家族非生物胁迫生物信息学分析[J]. 东北农业大学学报, 50(7):18-27.]

CAO C, MA K, 2016. Protein secondary structure assignment[J]. Chin J Bioinform, (14):187. [曹晨，马堃，2016. 蛋白质二级结构指定[J]. 生物信息学, (14):187.]

CHEN Q, 2011. The clon of a *AP2* transcription factor factor gene of maize and the construction of a prokaryotic expression vector[D]. Chengdu：Sichuan Agricultural University. [陈茜，2011. 一个玉米*AP2*类转录因子基因的克隆和原核表达载体构建[D]. 成都：四川农业大学.]

GUI XH, 2017. Culture conditions for oil accumulation and regulation of lipid synthetic metabolism of *chlorella protothecoides*[D]. Wuhan：Huazhong University of Science and Technology. [桂小华，2017. 原始小球藻产油条件及油脂合成代谢调控的研究[D]. 武汉：华中科技大学.]

GUO H, JIN SY, LIU H, et al., 2017. In silico cloning and bioinformatics analysis of *AP2/ERF* gene family from *Brassica oleracea*[J]. China Pharm, 20(1): 6-10. [郭慧，金司阳，刘寒，等，2017. 甘蓝*AP2/ERF*转录因子的克隆和生物信息学分析[J]. 中国药师，20(1)：6-10.]

HAN CZ, WANG J, 2016. Bioinformatics analysis of type VI secreted protein in *Xanthomonas Campestris* pv. *raphani* 756C[J]. J Huazhong Agric Univ, 35(4): 42-48. [韩长志，王娟，2016. 黄单胞菌*Xanthomonas campestris* pv. *raphani* 756C中VI型分泌蛋白的生物信息学分析[J]. 华中农业大学学报，35(4)：42-48.]

HOU LY, 2016. Isolation of cellulose-degrading strain *Pseudoxanthomonas* sp. J1 and its bioinformatics analysis[D]. Nanjing：Nanjing Agricultural University . [侯丽媛，2016. 纤维素降解菌*Pseudoxanthomonas* sp. J1的分离筛选及其生物信息学分析[D]. 南京：南京农业大学.]

JI Q, ZHOU F, ZHOU J, et al., 2018. Whole genome identification and bioinformatics analysis of *AP2/EREBP* transcription factors of *Ziziphus jujube*[J]. Genom Appl Biol, 37(7): 2983-2997. [纪晴，周凡，周军，等，2018. 枣*AP2/EREBP*转录因子的全基因组鉴定及生物信息学分析[J]. 基因组学与应用生物学，37(7)：2983-2997.]

JOFUKU KD, DEN BOER BG, VAN MONTAGU M, et al., 1994. Control of *Arabidopsis* flower and seed development by the homeotic gene *APETALA2*[J]. Plant Cell, 6(9): 1211-1225.

LI JF, GUO MZ, LIU Y, et al., 2008. An evolutionary tree reconstruction method combining quartet puzzling and neighbor joining[J]. J Comput Res Dev, 45(11): 1965-1973. [李建伏，郭茂祖，刘扬，等，2008. 一种基于QuartetPuzzling和邻接法的进化树构建算法[J]. 计算机研究与发展，45(11)：1965-1973.]

LIU Y, 2009. Evolutionary dynamics of flower development regulatory network[D]. Beijing：University of Chinese Academy of Sciences, University of Chinese Academy of Sciences. [刘杨，2009. 花发育调控网络的进化动态研究[D]. 北京：中国科学院研究生院;中国科学院大学.]

LIU Y, XIE LY, LAI ZX, et al., 2017. Cloning and bioinformatics analysis of *amaAG* in *Amaranthus tricolor* L[J]. Acta Agric Univ Jiangxi, 39(1):168-174. [柳燕，谢礼洋，赖钟雄，等，2017. 苋菜*amaAG*基因克隆与生物信息学分析[J]. 江西农业大学学报，

39(1) :168-174.]

MA HY, ZHAO QF, CHEN S, et al., 2020. Bioinformatics analysis of *PEPC* gene family in Pineapple[J].Chin J Trop Crop, 41(1)：97-103. [马海洋，赵秋芳，陈曙，等，2020. 菠萝*PEPC*基因家族生物信息学分析[J]. 热带作物学报, 41(1)：97-103.]

MA Q, LI JB, RUAN CJ, et al., 2016. Bioinformatics analysis of *WRI1* gene in *Hippophae rhamnoides*[J]. Hubei Agric Sci, 55(22). [马倩，李景滨，阮成江，等，2016. 沙棘*WRI1*转录因子基因的生物信息学分析[J]. 湖北农业科学, 55(22).]

MO XT, ZHAO J, FAN,YL, et al., 2013. Research progress on structure and function of maize transcription factors[J]. J Agric Sci Tech China, 15(3): 7-17. [莫晓婷，赵军，范云六，等，2013. 玉米转录因子结构与功能研究进展[J]. 中国农业科技导报，15(3)：7-17.]

NAKANO T, 2006. Genome-wide analysis of the *ERF* gene family in *Arabidopsis* and rice[J]. Plant Physiol, 140(2):411-432.

RICCARDO V, ANDREY Z, JASON A, et al., 2010. The genome of the domesticated apple (Malus [times] domestica Borkh.)[J]. Nat Genet, 42(10):833-839.

RUAN XL, ZHANG J, ZHANG, FL, et al., 2017. The bioinformatics analysis of AP2 gene family in Sweet Potato[J]. Mol Plant Breed, 15(06): 2066- 2072. [阮先乐，张杰，张福丽，等，2017. 甘薯AP2基因家族的生物信息学分析[J]. 分子植物育种，15(06)：2066- 2072.]

SCHWEDE T, KOPP J, GUEX N, et al., 2003. SWISS-MODEL: an automated protein homology- modeling server[J]. Nucl Acid Res, 31(13): 3381-3385.

SHI Y, XU,X, LI HY, et al., 2013. Bioinformatics analysis of the expansin gene family in rice[J]. Hereditas, 36(8): 809-820. [施杨，徐筱，李昊阳，等，2013. 水稻扩展蛋白家族的生物信息学分析[J]. 遗传，36(8)：809-820.]

SONG W, LI DL, WANG R, et al., 2013. Bioinformatics analysis of DELLA proteins in Rosaceous Plants[J]. Chin Agric Sci B, , 29(19):142-148. [宋伟，李鼎立，王然，等，2013. 蔷薇科植物DELLA蛋白的生物信息学分析[J]. 中国农学通报, 29(19):142-148.]

SU L, TONG SM, 2016. Sequence analysis and translation of AP2 / ERF transcription factor gene family in *Vitis vinifera*[J]. Biotechnol W, (4): 18-19. [栗丽，佟少明，2016. 葡萄(Vitis vinifera)AP2/ERF转录因子基因家族序列分析[J]. 生物技术世界，(4)：18-19.]

TAMURA K , PETERSON D , PETERSON N , et al., 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods[J]. Mol Biol Evol, 28(10):2731-2739.

TANG MF, 2007. The expression pattern of the *APETALA2* group genes in rice[D]. Yangling：NW A & F University. [唐美芳，2007. 水稻*APETALA2*基因组基因的表达特性分析[D]. 杨凌：西北农林科技大学.]

TU JQ, GAN L, FENG LL, et al., 2018. Bioinformatics analysis of *PEPC* gene family in *Arachis duranensis*[J]. J Trop Subtrop Bot, 26(2): 107-115. [涂嘉琦，甘璐，冯兰兰，等，2018. 蔓花生PEPC基因家族的生物信息学分析[J]. 热带亚热带植物学报，26(2): 107-115.]

WANG GH, GUI XH, DENG W, et al. 2013. Effect of "heterotrophic-stress"segmental culture on biomass and lipid content of *Chlorella protothecoides*[J]. China Biotechnol, 33(3): 99-104. [汪桂林，桂小华，邓伟，等，2013. "异养-胁迫"分段培养对原始小球藻生物量和油脂含量影响研究[J]. 中国生物工程杂志，33(3)：99-104.]

WANG GL, 2013. Two-step cultivation mode for *Chlorella protothecoides* and extraction of high-valued product[D]. Wuhan：Huazhong University of Science and Technology. [汪桂林，2013. 原始小球藻分段培养条件优化及高附加值产物提取工艺研究[D]. 武汉：华中科

技大学.]

WANG JS, LI LL, YANG H, et al., 2013. Electronic clone and bioinformatics analysis of F-box protein gene from wheat(*Triticum aestivum* L.)[J] J Henan Agric Sci，42(2):15-19. [王俊生，李俐俐，杨欢，等，2013. 小麦F-box蛋白基因的电子克隆和生物信息学分析[J]. 河南农业科学, 42(2):15-19.]

WANG LL, 2016. Genome-wide investigation of *AP2* subfamily in *Brassica napus* L. and *WRI1* in plants[D]. Yangling：Northwest A & F University. [王莉莉，2016. 甘蓝型油菜*AP2*亚家族转录因子及植物中*WRI1*的基因组学分析[D]. 杨凌：西北农林科技大学.]

WANG S, LI Y, LI MJ, et al., 2019. Cloning and bioinformatics analysis of *APs* gene in *Ricinus communis*[J]. Mol Plant Breed，17（22）：7344-7349. [王双，李跃，李孟建，等，2019. 蓖麻APs基因克隆及生物信息学分析[J]. 分子植物育种，17（22）：7344-7349.]

WANG Y, 2017. Stdy on expression and function of *AP2/ERF* transcription factor in flower bud of chinese cherry during dormancy release[D]. Jinhua：Zhejiang Normal University. [王月，2017. 中国樱桃*AP2/ERF*转录因子在花芽休眠解除过程的表达与作用研究[D]. 金华：浙江师范大学.]

WEN KX, LIU XM, 2010. The important role of *AP2* functional genes in plant floral development[J]. Biotechnol Bull, (2): 1-7. [闻可心，刘雪梅，2010. *AP2*功能基因在植物花发育中的重要作用[J]. 生物技术通报，(2)：1-7.]

WESSLER SR . 2005. Homing into the origin of the AP2 DNA binding domain[J]. Trends Plant Sci,34(08): 7-13.

WU LJ, YU ZL, LIN XZ, et al., 2016. Advances and perspectives of genetic engineering in *Chlorella*[J]. Bull Sci Technol, 32(12):51-56. [吴丽娟，余志良，林祥志，等，2016. 小球藻基因工程的研究进展及应用前景展望[J]. 科技通报，32(12) :51-56.]

WU ZJ, LU L, LI XH, et al., 2014. Isolation and expression profiles analysis of *AP2/ERF-B3* group transcription factor from *Camellia sinensis*[J]. J Nanjing Agric Univ, 37(4): 67-75. [吴致君，卢莉，黎星辉，等，2014. 茶树*AP2/ERF-B3*类转录因子基因的克隆与表达特性分析[J]. 南京农业大学学报，37(4)：67-75.]

XIAO J, WU XY, WANG LL, et al., 2011. Bioinformatics of profilin protein from peanut[J]. Immunol J, 27(2):158-161. [肖杰，吴序栎，王琳琳，等，2011. 花生profilin蛋白的生物信息学分析[J]. 免疫学杂志, 27(2):158-161.]

XIA JL, WAN MX, WANG RM, et al., 2009. Current situation and progress of Microalgae Biodiesel[J]. China Biotechnol, 29(7): 118~126. [夏金兰，万民熙，王润民，等，2009. 微藻生物柴油的现状与进展[J]. 中国生物工程杂志，29(7)：118-126.]

XIA QY, 2007. Cloning and Expression of venom allergen *soli1* and *soli4* genes in the *Solenopsis invicta*[D]. Chongqing：Southwest University. [夏巧玉，2007. 红火蚁毒素致敏原*soli1*、*soli4*基因的克隆及表达[D]. 重庆：西南大学.]

XU ZJ, LIU Y, XU L, et al., 2019. Bioinformatics analysis of *NF-YB* gene family of transcription factor in *Zea mays*[J]. Mol Plant Breed,17(12):3807-3816. [徐志军，刘洋，徐磊，等，2019. 玉米转录因子*NF-YB*基因家族的生物信息学分析[J]. 分子植物育种，17(12):3807-3816.]

YANG L, 2015. Clone and functional analysis of carbonic anhydrase gene of tobacco[D]. Chengdu：Sichuan Agricultural University. [杨朗，2015. 烟草碳酸酐酶基因的克隆及其生物信息学分析[D]. 成都：四川农业大学.]

ZHANG G, CHEN M, CHEN X, et al., 2008. Phylogeny, gene structures, and expression patterns of the *ERF* gene family in soybean (*Glycine max* L.)[J]. J Exp Bot, 59(15):4095-4107.

ZHANG HY, LI ZT, ZHAO CJ, et al., 2013. Genome-wide analysis of LIM domain-containing protein gene family in maize[J]. J Maize Sci, (3): 40-47. [张海燕，李佐同，赵长江，等，2013. 玉米LIM结构域蛋白基因家族分析[J]. 玉米科学，(3)：40-47.]

ZHANG JY, WANG QJ, GUO ZR, 2012. Progresses on plant *AP2/ERF* transcription factors[J]. Hereditas, 34(7): 835-847. [张计育，王庆菊，郭忠仁，2012. 植物*AP2/ERF*类转录因子研究进展[J]. 遗传，34(7)：835-847.]

ZHANG Q, CHENG J, LI L, et al., 2018. Research progress on plant *AP2/ERF* transcription factor family[J]. Biotechnol B, 34(8): 1-7. [张麒，陈静，李俐，等，2018. 植物*AP2/ERF*转录因子家族的研究进展[J]. 生物技术通报，34(8)：1-7.]

ZHANG Y, SUN FB, 2011. Function of *AP2* in plant flower development[J]. China's Sci Technol,(23): 436-436. [张妍，孙丰宾，2011. *AP2*在植物花发育中的功能简述[J]. 中国科技纵横，(23)：436-436.]

ZHAO JL, YAO WJ, WANG SJ, et al., 2015. *AP2/ERF* gene family in *Populus trichocarpa* by bioinformatics[J]. J NE For Univ, 43(10): 21-29. [赵金玲，姚文静，王升级，等，2015. 杨树*AP2/ERF*转录因子家族生物信息学分析[J]. 东北林业大学学报，43(10)：21-29.]

ZHAO LF, CHAI TY, 2008. Role of *AP2/EREBP* transcription factors in plant development and stress response[J]. Chin Bull Bot, 25(1): 89-101. [赵利锋，柴团耀，2008. *AP2/EREBP*转录因子在植物发育和胁迫应答中的作用[J]. 植物学通报，25(1)：89-101.]

ZHU CS, LIU Y, CHENG MM, et al., 2018. Cloning and expression analysis of two *AP2/ERF* transcription factors in *Tripterygium wilfordii* Hook.f[J]. Acta Bot Boreal-Occident Sin, 38(8): 43-50. [祝传书，刘艳，陈蒙蒙，等，2018. 雷公藤*AP2/ERF*转录因子基因的克隆与分析[J]. 西北植物学报，38(8)：43-50.]

ZHU GZ, 2016. Bioinformatics analysis of *AP2/ERF* gene family in *phalaenopsis equestris*[D]. Beijing: Chinese Academy of Forestry. [朱光哲，2016. 小兰屿蝴蝶兰AP2/ERF基因家族生物信息学分析[D]. 北京：中国林业科学研究院.]

ZHU YM, YU JY, YU Y, et al., 2019. *GsRAV3* gene, encoding a member of the *AP2/EREBP* family from *Glycine soja*, negatively regulates plant ABA sensitivity in *Arabidopsis*[J]. J NE Agric Univ, 50(5): 8-18. [朱延明，于纪洋，于洋，等，2019. 野生大豆AP2/RAV亚家族转录因子*GsRAV3*负调控拟南芥对ABA的敏感性[J]. 东北农业大学学报，(5)：8-18.]

ZHUANG J, PENG RH , CHENG ZM , et al., 2008. Genome-wide analysis of the putative *AP2/ERF* family genes in *Vitis vinifera*[J]. Sci Hort (Amsterdam), 123(1): 1-81.

ZHUANG J, ZHOU XR, SUN CC, et al., Cloning and bioinformatics analysis of a class of *AP2/ERF* transcription factors in *Brassica napus* L.[J]. China Biotechnol, 28(5): 29-40. [庄静，周熙荣，孙超才，等，2008. 甘蓝型油菜中一类*AP2/ERF*转录因子的克隆和生物信息学分析[J]. 中国生物工程杂志，28(5): 29-40.]